



(12) 发明专利

(10) 授权公告号 CN 106919465 B

(45) 授权公告日 2021.03.16

(21) 申请号 201510983346.2

(22) 申请日 2015.12.24

(65) 同一申请的已公布的文献号
申请公布号 CN 106919465 A

(43) 申请公布日 2017.07.04

(73) 专利权人 伊姆西IP控股有限责任公司
地址 美国马萨诸塞州

(72) 发明人 骆渊 徐宁 马思超 商秋
梁韶聪

(74) 专利代理机构 北京市金杜律师事务所
11256

代理人 王茂华

(51) Int. Cl.
G06F 11/14 (2006.01)

(56) 对比文件

US 7225307 B2, 2007.05.29

CN 103814360 A, 2014.05.21

CN 101030154 A, 2007.09.05

审查员 唐进岭

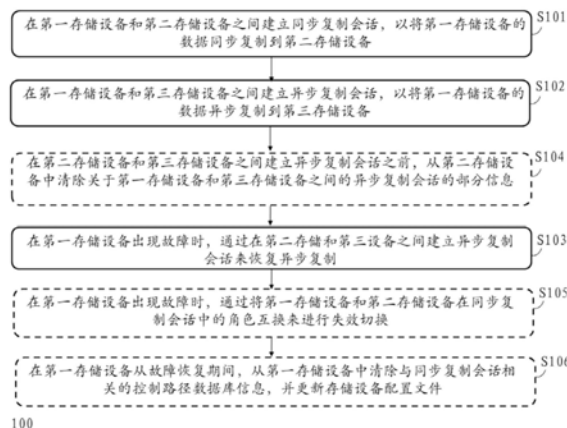
权利要求书3页 说明书8页 附图2页

(54) 发明名称

用于存储系统中多重数据保护的方法和装置

(57) 摘要

本公开的实施例提出了用于存储系统中的多重数据保护的方法和装置。该存储系统包括第一存储设备、第二存储设备和第三存储设备，该方法包括：在该第一存储设备和该第二存储设备之间建立同步复制会话，以将该第一存储设备的数据同步复制到该第二存储设备；在该第一存储设备和该第三存储设备之间建立异步复制会话，以将该第一存储设备的数据异步复制到该第三存储设备；以及在该第一存储设备出现故障时，通过在该第二存储和该第三设备之间建立异步复制会话来恢复该异步复制。根据本公开的实施例的方法或装置，能够同时利用同步复制和异步复制技术来提供本地和远程多重数据保护。



1. 一种用于存储系统中的多重数据保护的方法,所述存储系统包括第一存储设备、第二存储设备和第三存储设备,所述方法包括:

在所述第一存储设备和所述第二存储设备之间建立同步复制会话,以将所述第一存储设备的数据同步复制到所述第二存储设备;

在所述第一存储设备和所述第三存储设备之间建立异步复制会话,以将所述第一存储设备的数据异步复制到所述第三存储设备;以及

在所述第一存储设备出现故障时,通过在所述第二存储和所述第三设备之间建立异步复制会话来恢复所述异步复制,

在所述第一存储设备出现故障时,通过在所述同步复制会话中将所述第二存储设备的数据同步复制到所述第一存储设备来进行失效切换;以及

在所述第一存储设备从故障恢复期间,从所述第一存储设备中清除与所述同步复制会话相关的控制路径数据库信息,并更新存储设备配置文件。

2. 根据权利要求1所述的方法,其中

将所述第一存储设备的数据同步复制到所述第二存储设备包括:将所述第一存储设备中的、关于所述第一存储设备和所述第三存储设备之间的异步复制会话的信息同步到所述第二存储设备;并且

所述方法进一步包括:在所述第二存储设备和所述第三存储设备之间建立异步复制会话之前,从所述第二存储设备中清除关于所述第一存储设备和所述第三存储设备之间的异步复制会话的部分信息。

3. 根据权利要求2所述的方法,其中从所述第二存储设备中清除关于所述第一存储设备和所述第三存储设备之间的异步复制会话的信息包括:

从同步复制到所述第二存储设备的数据中识别所述第一存储设备和所述第三设备之间的异步复制会话的会话签名,以及

从所述第二存储设备中清除由所述会话签名所指示的异步复制会话的信息。

4. 根据权利要求3所述的方法,其中

所述会话签名包括会话标识符和文件系统标识符,并且

从所述第二存储设备中清除由所述会话签名所指示的异步复制会话的信息包括:

在所述会话签名所包括的文件系统标识符与从所述第一存储设备同步复制到所述第二存储设备中的控制路径数据库中的内部文件系统检查点标识符匹配时,从所述第二存储设备中清除由所述会话标识符所指示的异步复制会话的信息。

5. 根据权利要求2所述的方法,其中将所述第一存储设备中的、关于所述第一存储设备和所述第三存储设备之间的异步复制会话的信息同步到所述第二存储设备包括:

将所述第一存储设备中的、关于所述第一存储设备和所述第三存储设备之间的异步复制会话的信息作为数据的一部分同步到所述第二存储设备;或者

将所述第一存储设备中的、关于所述第一存储设备和所述第三存储设备之间的异步复制会话的信息作为内部文件系统检查点的一部分同步到所述第二存储设备。

6. 根据权利要求1所述的方法,其中在所述第二存储设备和所述第三设备之间建立异步复制会话包括:

向所述第二存储设备加载文件系统;以及

在所述第二存储设备和所述第三存储设备之间建立文件系统级的异步复制会话。

7. 根据权利要求1所述的方法,其中更新存储设备的配置文件包括:修改原有的配置文件中的文件系统加载项。

8. 一种用于存储系统中的多重数据保护的装置,所述存储系统包括第一存储设备、第二存储设备和第三存储设备,所述装置包括:

同步复制单元,被配置为在所述第一存储设备和所述第二存储设备之间建立同步复制会话,以将所述第一存储设备的数据同步复制到所述第二存储设备;

异步复制单元,被配置为在所述第一存储设备和所述第三存储设备之间建立异步复制会话,以将所述第一存储设备的数据异步复制到所述第三存储设备;异步复制恢复单元,被配置为在所述第一存储设备出现故障时,通过在所述第二存储和所述第三设备之间建立异步复制会话来恢复所述异步复制;

失效切换单元,被配置为在所述第一存储设备出现故障时,通过在所述同步复制会话中将所述第二存储设备的数据同步复制到所述第一存储设备来进行失效切换;以及

文件更新单元,被配置为在所述第一存储设备从故障恢复期间,从所述第一存储设备中清除与所述同步复制会话相关的控制路径数据库信息,并更新存储设备配置文件。

9. 根据权利要求8所述的装置,其中所述同步复制单元进一步被配置为将所述第一存储设备中的、关于所述第一存储设备和所述第三存储设备之间的异步复制会话的信息也同步到所述第二存储设备;并且

所述装置进一步包括:

第一清除单元,被配置为在所述第二存储设备和所述第三存储设备之间建立异步复制会话之前,从所述第二存储设备中清除关于所述第一存储设备和所述第三存储设备之间的异步复制会话的部分信息。

10. 根据权利要求9所述的装置,其中所述第一清除单元进一步包括:

识别单元,被配置为从同步复制到所述第二存储设备的数据中识别所述第一存储设备和所述第三设备之间的异步复制会话的会话签名,以及

信息清除单元,被配置为从所述第二存储设备中清除由所述会话签名所指示的异步复制会话的信息。

11. 根据权利要求10所述的装置,其中所述会话签名包括会话标识符和文件系统标识符,并且所述信息清除单元进一步被配置为:

在所述会话签名所包括的文件系统标识符与从所述第一存储设备同步复制到所述第二存储设备中的控制路径数据库中的内部文件系统检查点标识符匹配时,从所述第二存储设备中清除由所述会话标识符所指示的异步复制会话的信息。

12. 根据权利要求9所述的装置,其中,所述同步复制单元进一步被配置为通过以下方式之一将所述第一存储设备中的、关于所述第一存储设备和所述第三存储设备之间的异步复制会话的信息也同步到所述第二存储设备:

将所述第一存储设备中的、关于所述第一存储设备和所述第三存储设备之间的异步复制会话的信息作为数据的一部分同步到所述第二存储设备;或者

将所述第一存储设备中的、关于所述第一存储设备和所述第三存储设备之间的异步复制会话的信息作为内部文件系统检查点的一部分同步到所述第二存储设备。

13. 根据权利要求8所述的装置,其中所述异步复制单元进一步被配置为通过以下方式在所述第二存储设备和所述第三设备之间建立异步复制会话:

向所述第二存储设备加载文件系统;以及

在所述第二存储设备和所述第三存储设备之间建立文件系统级的异步复制会话。

14. 根据权利要求8所述的装置,其中所述文件更新单元被配置为通过修改原有的配置文件中的文件系统加载项来更新存储设备的配置文件。

15. 一种用于存储系统中多重数据保护的装置,所述装置包括至少一个处理器;以及包括计算机程序代码的至少一个存储器,其中所述至少一个存储器和所述计算机程序代码被配置为:与所述至少一个处理器一起,促使所述装置执行根据权利要求1-7中任一权利要求所述的方法。

用于存储系统中多重数据保护的方法和装置

技术领域

[0001] 本公开的实施例涉及数据存储系统,并且更具体地涉及用于多重数据保护的方法和装置。

背景技术

[0002] 在数据存储系统中,为了提高数据的安全性和可用性,可以同时存储数据的多个备份。例如,可以通过镜像视图技术将数据从一个存储设备复制到另一存储设备。目前数据复制可以通过同步复制或者异步复制技术来进行。

[0003] 同步复制是网络附属存储(NAS)在本地局域网(LAN)或广域网络(WAN)上通过块(Block)复制技术而产生多个数据同步复件的技术。同步复制的优点是数据同步保护时延小,恢复点目标(RPO)等于0。因此当生产数据中心发生灾难事件时,不会造成数据丢失;不足之处在于只在相对短距离(例如10Km)内工作。因此,同步复制适用于恢复时间目标(RTO)低的应用。

[0004] 相对于同步复制而言,在异步复制技术中,在将数据复制到辅站点之前存在更长的时延。异步复制能够支持更远的距离,例如100Km。

[0005] 因此,同步复制和异步复制各有优缺点,适用于不同的场景。在目前的存储系统中,还没有支持同步复制和异步复制共存的技术方案。

发明内容

[0006] 为了改善数据保护,本公开的实施例提出了用于存储系统中的数据保护的方法和装置。

[0007] 下面给出了对各实施例的简要概述,以提供对各种实施例的一些方面的基本理解。该概述不旨在标识关键元素的要点或描述各种实施例的范围。其唯一目的在于以简化形式呈现一些概念,作为对后述更具体描述的前序。

[0008] 本公开的第一方面提供了一种用于存储系统中的多重数据保护的方法,该存储系统包括第一存储设备、第二存储设备和第三存储设备,该方法包括:在该第一存储设备和该第二存储设备之间建立同步复制会话,以将该第一存储设备的数据同步复制到该第二存储设备;在该第一存储设备和该第三存储设备之间建立异步复制会话,以将该第一存储设备的数据异步复制到该第三存储设备;以及在该第一存储设备出现故障时,通过在该第二存储和该第三设备之间建立异步复制会话来恢复该异步复制。

[0009] 在一个实施例中,将该第一存储设备的数据同步复制到该第二存储设备可以包括:将该第一存储设备中的、关于该第一存储设备和该第三存储设备之间的异步复制会话的信息也同步到该第二存储设备;并且该方法进一步包括:在该第二存储设备和该第三存储设备之间建立异步复制会话之前,从该第二存储设备中清除关于该第一存储设备和该第三存储设备之间的异步复制会话的部分信息。

[0010] 在另一实施例中,从该第二存储设备中清除关于该第一存储设备和该第三存储设

备之间的异步复制会话的信息可以包括:从同步复制到该第二存储设备的数据中识别该第一存储设备和该第三设备之间的异步复制会话的会话签名,以及从该第二存储设备中清除由该会话签名所指示的异步复制会话的信息。在又一实施例中,该会话签名可以包括会话标识符和文件系统标识符,并且从该第二存储设备中清除由该会话签名所指示的异步复制会话的信息可以包括:在该会话签名所包括的文件系统标识符与从该第一存储设备导入到该第二存储设备中的控制路径数据库中的内部文件系统检查点标识符匹配时,从该第二存储设备中清除由该会话标识符所指示的异步复制会话的信息。

[0011] 在一个实施例中,将该第一存储设备中的、关于该第一存储设备和该第三存储设备之间的异步复制会话的信息也同步到该第二存储设备可以包括:将该第一存储设备中的、关于该第一存储设备和该第三存储设备之间的异步复制会话的信息作为数据的一部分同步到该第二存储设备;或者将该第一存储设备中的、关于该第一存储设备和该第三存储设备之间的异步复制会话的信息作为内部文件系统检查点的一部分同步到该第二存储设备。

[0012] 在另一实施例中,在该第二存储设备和该第三设备之间建立异步复制会话可以包括:向该第二存储设备加载文件系统;以及在该第二存储设备和该第三存储设备之间建立文件系统级的异步复制会话。

[0013] 在一个实施例中,该方法可以进一步包括:在该第一存储设备出现故障时,通过将该第一存储设备和该第二存储设备在该同步复制会话中的角色互换来进行失效切换;以及在在该第一存储设备从故障恢复期间,从该第一存储设备中清除与该同步复制会话相关的控制路径数据库信息,并更新存储设备配置文件以保证该第一存储设备重启后的文件系统加载过程不因文件系统的不匹配而导致系统循环重启。在另一实施例中,更新存储设备的配置文件可以包括:修改原有的配置文件中的文件系统加载项。

[0014] 本公开的第二方面提供了一种用于存储系统中的多重数据保护的装置,该存储系统包括第一存储设备、第二存储设备和第三存储设备,该装置包括:同步复制单元,被配置为在该第一存储设备和该第二存储设备之间建立同步复制会话,以将该第一存储设备的数据同步复制到该第二存储设备;异步复制单元,被配置为在该第一存储设备和该第三存储设备之间建立异步复制会话,以将该第一存储设备的数据异步复制到该第三存储设备;以及异步复制恢复单元,被配置为在该第一存储设备出现故障时,通过在该第二存储和该第三设备之间建立异步复制会话来恢复该异步复制。

[0015] 在一个实施例中,同步复制单元可以进一步被配置为将该第一存储设备中的、关于该第一存储设备和该第三存储设备之间的异步复制会话的信息也同步到该第二存储设备;并且该装置可以进一步包括:第一清除单元,被配置为在该第二存储设备和该第三存储设备之间建立异步复制会话之前,从该第二存储设备中清除关于该第一存储设备和该第三存储设备之间的异步复制会话的部分信息。

[0016] 在另一实施例中,第一清除单元可以进一步包括:识别单元,被配置为从同步复制到该第二存储设备的数据中识别该第一存储设备和该第三设备之间的异步复制会话的会话签名,以及信息清除单元,被配置为从该第二存储设备中清除由该会话签名所指示的异步复制会话的信息。在又一实施例中,会话签名可以包括会话标识符和文件系统标识符,并且该信息清除单元可以进一步被配置为:在该会话签名所包括的文件系统标识符与从该第

一存储设备导入到该第二存储设备中的控制路径数据库中的内部文件系统检查点标识符匹配时,从该第二存储设备中清除由该会话标识符所指示的异步复制会话的信息。

[0017] 在另一实施例中,该同步复制单元进一步被配置为通过以下方式之一将该第一存储设备中的、关于该第一存储设备和该第三存储设备之间的异步复制会话的信息也同步到该第二存储设备:将该第一存储设备中的、关于该第一存储设备和该第三存储设备之间的异步复制会话的信息作为数据的一部分同步到该第二存储设备;或者将该第一存储设备中的、关于该第一存储设备和该第三存储设备之间的异步复制会话的信息作为内部文件系统检查点的一部分同步到该第二存储设备。

[0018] 在又一实施例中,异步复制单元进一步被配置为通过以下方式在该第二存储设备和该第三设备之间建立异步复制会话:向该第二存储设备加载文件系统;以及在该第二存储设备和该第三存储设备之间建立文件系统级的异步复制会话。

[0019] 在一个实施例中,该装置可以进一步包括:失效切换单元,被配置为在该第一存储设备出现故障时,通过将该第一存储设备和该第二存储设备在该同步复制会话中的角色互换来进行失效切换;以及文件更新单元,被配置为在该第一存储设备从故障恢复期间,从该第一存储设备中清除与该同步复制会话相关的控制路径数据库信息,并更新存储设备配置文件以保证该第一存储设备重启后的文件系统加载过程不因文件系统的不匹配而导致系统循环重启。在另一实施例中,文件更新单元可以被配置为通过修改原有的配置文件中的文件系统加载项来更新存储设备的配置文件。

[0020] 本公开的第三方面提供了一种装置,该装置包括至少一个处理器;以及包括计算机程序代码的至少一个存储器,其中该至少一个存储器和该计算机程序代码被配置为:与该至少一个处理器一起,促使该装置执行根据本公开的第一方面的方法。

[0021] 根据本公开的实施例的方法或装置,能够同时利用同步复制和异步复制技术来提供本地和远程多重数据保护。

[0022] 尽管在附图中通过示例的方式示出了特定的实施例,然而,应当理解,本文的具体实施例的描述不意在将实施例限制为所公开的具体形式。

附图说明

[0023] 从下文的公开内容和权利要求中,本公开的目的、优点和其他特征将变得更加明显。这里仅出于示例的目的,参考附图来给出优选实施例的非限制性描述,在附图中:

[0024] 图1示出了根据本公开的实施例的用于存储系统中的多重数据保护的方法的流程图;

[0025] 图2a-图2b示出了根据本公开的实施例的在第一存储设备故障前和故障期间数据保护的示意图;以及

[0026] 图3示出了根据本公开的实施例的装置的示例性结构图。

具体实施方式

[0027] 在以下描述中,出于说明的目的而阐述许多细节。然而,本领域普通技术人员将认识到可以在不使用这些具体细节的情况下实现本公开的实施例。因此,本公开不旨在于受限于所示实施例、而是将被赋予与本文描述的原理和特征一致的最宽的范围。

[0028] 应当理解,术语“第一”、“第二”等仅被用来将一个元素与另一个元素区分开来。而实际上第一元素也能够被称为第二元素,反之亦然。另外还应当理解“包括”,“包含”仅被用来说明所陈述的特征、元素、功能或者部件的存在,然而并不排除存在一个或者多个其他的特征、元素、功能或者部件。

[0029] 为了便于解释,本文的一些实施例将以EMC™公司的VNX2产品为例来介绍相应的方法或装置,然而,如本领域技术人员可以理解的,本公开的实施例绝不限于应用于VNX2产品,而是可以被应用于任何存在类似问题的存储系统中。

[0030] 如前所述,在现有技术中并不存在使同步复制和异步复制共存的解决方案。例如,在VNX2存储产品中,可以提供复制器V2(简称为Repv2)用于异步复制,或者提供虚拟数据移动器(简称为VDM)同步用于同步复制。

[0031] 为使同步复制和异步复制共存、以同时利用两者的优点,将面临诸多问题,例如,如何使同步复制和异步复制在存储系统中协作、如何在发生故障时恢复会话等。

[0032] 为了解决以上问题中的至少一些问题,本公开的实施例提出了用于存储系统中的多重数据保护的方法。图1中示出了该方法100的示例性流程图。该存储系统包括第一存储设备、第二存储设备、和第三存储设备。其中该存储系统可以是,但不限于,VNX2存储系统,并且其中的存储设备可以是(但不限于)存储阵列。

[0033] 如图1所示,该方法100包括,在S101,在第一存储设备和第二存储设备之间建立同步复制会话,以将第一存储设备的数据同步复制到第二存储设备;这意味着第一存储设备的数据能够以低延时复制到第二存储设备。在S102,在第一存储设备和第三存储设备之间建立异步复制会话,以将第一存储设备的数据异步复制到第三存储设备;由于异步复制的较大时延,这意味着第三存储设备中保存的第一存储设备的信息可能不是最新的。在S103,在第一存储设备出现故障时,通过在第二存储和第三设备之间建立异步复制会话来恢复异步复制。

[0034] 图2a-图2b中示意性地示出利用该方法100的存储系统在故障前和故障期间数据保护的示意图。

[0035] 通过该方法,提供了存储系统中的本地和远程双重数据保护,例如文件保护。如本领域技术人员能够理解的,本方法还可以应用于具有更多存储设备的存储系统。通过该方法,提供了同步复制和异步复制协作的机制,并且,使得在第一存储设备发生故障时,能够将同步地保存在第二存储设备中的、第一设备的数据异步复制到第三存储设备。

[0036] 在一个实施例中,在块S103中,在第一存储设备出现故障时,在该第二存储设备和该第三设备之间建立异步复制会话可以包括:向第二存储设备加载文件系统;以及在第二存储设备和第三存储设备之间建立文件系统级的异步复制会话。

[0037] 目前还没有如方法100该的同时利用同步复制和异步复制的解决方案。以传统VNX2产品为例,由于以下原因中的至少一些原因,其目前不能够支持同步复制和异步复制的共存:

[0038] 1. 传统的VNX2中,可同步复制的VDM不能够创建文件系统(FS)级的异步复制Repv2会话。如果在存储设备A和存储设备B之间建立了同步复制,同时在存储设备A和存储设备C之间建立了异步复制,则在存储设备A发生故障时,同步复制将从存储设备A向存储设备B进行失效切换。这导致VDM的文件系统数据被同步到存储设备B侧,并且网络附属存储数据库

(NASDB) 将从NASDB副本控制逻辑单元号 (Control LUN) 导入。然而由于在存储设备B的数据路径数据库 (例如, 保存卷) 中存在来自存储设备A的关于设备A和设备C之间的异步复制会话 (Repv2) 的信息, 导致不能够在存储设备B的VDM中加载文件系统。

[0039] 2. 这种情况下, 在存储设备B中的、从存储设备A中同步来的FS Repv2是断裂的, 而且该Repv2无法通过传统的nas_replicate命令来删除, 因为同步复制特征 (Syncrep) 仅将VDM数据失效切换到存储设备B, 该VDM数据包括VDM根文件系统和用户文件系统, 但是不包括操作系统DART的根文件系统。因此, 在存储设备B中, DART的根文件系统中的数据库NameDB中未存储Repv2会话信息。

[0040] 3. 在Syncrep失效切换发生后, 原有的源存储设备 (即, A) 从故障中恢复。这时, 由于失效切换后该存储设备A处于备用状态, 其LUN块处于只读状态, 在设备启动过程中装载LUN块上的文件系统将出现失败而导致设备循环重启。

[0041] 可选地, 为了利用现有的存储设备 (例如VNX2) 来实施本地和远程双重数据保护, 在一些实施例中, 在方法100的块S101, 可以将第一存储设备中的、关于第一存储设备和第三存储设备之间的异步复制会话的信息 (例如Repv2信息) 也同步到第二存储设备。在这些实施方式中, 该方法100可以进一步包括: 在块S104, 在该第二存储设备和该第三存储设备之间建立异步复制会话之前, 从该第二存储设备中清除关于该第一存储设备和该第三存储设备之间的异步复制会话的部分信息。该实施例能够强制删除第二存储设备中的、不利于在第二设备和第三设备之间建立异步会话的信息。

[0042] 在一个实施例中, 在块S101, 可以将第一存储设备中的、关于第一存储设备和第三存储设备之间的异步复制会话的信息作为数据的一部分同步到第二存储设备; 或者将作为内部文件系统检查点的一部分同步到第二存储设备。

[0043] 在另一实施例中, 在利用以上所述的VNX2存储产品来实施该方法时, 为了能够在第二存储设备的VDM中加载文件系统, 可以在块S104中, 从第二设备中清除关于第一和第三存储设备之间的异步复制会话的信息中影响文件加载的信息。而对于不影响文件系统加载的其余信息, 可以保留。强制删除的信息例如可以是关于第一存储设备和第三存储设备之间的断裂的Repv2的信息, 该信息通过传统的nas_replicate命令是无法删除的。

[0044] 如本领域技术人员能够理解的, 在S104中所清除的信息可以根据该所应用于的存储系统的不同而不同, 并且对于某些存储系统, 可以不需要该清除操作。

[0045] 在另一实施例中, 在块S104中从第二存储设备中清除关于第一存储设备和第三存储设备之间的异步复制会话的信息可以包括: 从同步复制到第二存储设备的数据中识别第一存储设备和第三设备之间的异步复制会话的会话签名, 以及从第二存储设备中清除由该会话签名所指示的异步复制会话的信息。

[0046] 在又一实施例中, 会话签名可以包括会话标识符 (Session ID) 和文件系统标识符 (FS ID), 并且, 在块S104中, 可以在该会话签名所包括的FS ID与从第一存储设备导入到第二存储设备中的控制路径数据库 (例如, 网络附属系统数据库NASDB) 中的内部文件系统检查点标识符 (CheckPoint FS ID) 匹配时, 从第二存储设备中清除由该会话标识符所指示的异步复制会话的信息。再例如, 可以通过将存储在第二存储设备的保存卷 (Save Volume) 中的信息与存储在第二存储设备中的控制站的NASDB中的信息进行匹配、计算正确的DART, 并且利用合适的参数使DART删除断裂的Repv2。在删除Repv2之后, 文件系统可以被加载到第

二存储设备的VDM中,以继续提供数据服务。

[0047] 如本领域技术人员能够理解的,在另外的实施例中,从第二存储设备中清除的信息也可以由其他标识所指示,例如连接ID,目标地址ID等。

[0048] 在一个实施例中,该方法100还可以包括,在块S105,在第一存储设备出现故障时,通过将第一存储设备和第二存储设备在同步复制会话中的角色互换来进行失效切换。例如,角色互换可以是在所述同步复制会话中将第二存储设备的数据同步复制到第一存储设备。作为示例,失效切换可以包括将第一存储设备由活动状态转为备用状态,同时将第二存储设备由备用状态转为活动状态。

[0049] 在块S106,在第一存储设备从故障恢复期间,从第一存储设备中清除与该同步复制会话相关的控制路径数据库信息,并更新存储设备配置文件。这有助于确保第一存储设备重启后的文件系统加载过程不会由于文件系统的不匹配而导致系统循环重启。在一个实施例中,可以通过修改原有的配置文件中的文件系统加载项来更新存储设备的配置文件。

[0050] 如上所述,在未应用本公开的方法100时,在第一存储设备失效切换发生后,一旦该第一存储设备从故障中恢复,则在第一存储设备重启后试图加载文件系统时,可能出现循环重启。这是因为失效切换后第一存储设备处于备用状态,其LUN块处于只读状态,但是其VDM中有加载的文件系统,因此重启后执行加载文件系统的写操作将触发重启。而利用本公开的方法100的一些实施例,可以通过在不进行与DART的互联的情况下,清除第一存储设备的控制站中与同步会话有关的文件系统NASDB表项,来解决该问题。控制站可以建立新的配置文件(例如,boot.cfg),例如通过修改原有的配置文件中的文件系统加载项来建立新的配置文件;从而在DART重启后,能够根据新的配置文件来加载文件系统,从而避免循环重启。

[0051] 以下参考附图3描述用于存储系统中的多重数据保护的装置300的结构。装置300所适用的存储系统包括多个存储设备,例如第一存储设备、第二存储设备和第三存储设备。例如该存储系统可以是(但是并不限于)VNX2系统。装置300可以执行参考图1所述的方法100,但是不限于执行该方法100。同样该方法100可以由装置300执行,但是不限于由装置300执行。例如方法100的至少一些操作可以由其它的装置来执行。

[0052] 在一个实施例中,该装置300包括同步复制单元301,被配置为在该第一存储设备和该第二存储设备之间建立同步复制会话,以将该第一存储设备的数据同步复制到该第二存储设备;异步复制单元302,被配置为在该第一存储设备和该第三存储设备之间建立异步复制会话,以将该第一存储设备的数据异步复制到该第三存储设备;以及异步复制恢复单元303,被配置为在该第一存储设备出现故障时,通过在该第二存储和该第三设备之间建立异步复制会话来恢复该异步复制。

[0053] 在一个实施例中,单元301-303可以被配置为执行参考图1所述的方法100的步骤S101-S103。因此,参考方法100进行的相关描述在此同样适用。

[0054] 在一个实施例中,同步复制单元301进一步被配置为将第一存储设备中的、关于第一存储设备和第三存储设备之间的异步复制会话的信息也同步到该第二存储设备;并且该装置300可以进一步包括:第一清除单元304,被配置为在该第二存储设备和该第三存储设备之间建立异步复制会话之前,从该第二存储设备中清除关于该第一存储设备和该第三存储设备之间的异步复制会话的部分信息。

[0055] 在一个实施例中,以上参考方法100的块S104的操作在此同样适用于第一清除单元304,因此不再赘述。

[0056] 在另一实施例中,第一清除单元可以进一步包括:识别单元3041,被配置为从同步复制到该第二存储设备的数据中识别该第一存储设备和该第三设备之间的异步复制会话的会话签名,以及信息清除单元3042,被配置为从该第二存储设备中清除由该会话签名所指示的异步复制会话的信息。

[0057] 在又一实施例中,会话签名可以包括会话标识符和文件系统标识符,并且信息清除单元3042可以进一步被配置为:在该会话签名所包括的文件系统标识符与从该第一存储设备导入到该第二存储设备中的控制路径数据库中的内部文件系统检查点标识符匹配时,从该第二存储设备中清除由该会话标识符所指示的异步复制会话的信息。

[0058] 作为一个示例,同步复制单元301可以进一步被配置为通过以下方式之一将该第一存储设备中的、关于该第一存储设备和该第三存储设备之间的异步复制会话的信息也同步到该第二存储设备:将该第一存储设备中的、关于该第一存储设备和该第三存储设备之间的异步复制会话的信息作为数据的一部分同步到该第二存储设备;或者将该第一存储设备中的、关于该第一存储设备和该第三存储设备之间的异步复制会话的信息作为内部文件系统检查点的一部分同步到该第二存储设备。

[0059] 作为另一示例,异步复制单元302可以进一步被配置为通过以下方式在该第二存储设备和该第三设备之间建立异步复制会话:向该第二存储设备加载文件系统;以及在该第二存储设备和该第三存储设备之间建立文件系统级的异步复制会话。

[0060] 附加地或者替代地,装置300可以包括失效切换单元305,被配置为在该第一存储设备出现故障时,通过将该第一存储设备和该第二存储设备在该同步复制会话中的角色互换来进行失效切换;以及文件更新单元306,被配置为在该第一存储设备从故障恢复期间,从该第一存储设备中清除与该同步复制会话相关的控制路径数据库信息,并更新存储设备配置文件以保证该第一存储设备重启后的文件系统加载过程不因文件系统的不匹配而导致系统循环重启。

[0061] 在一个实施例中,文件更新单元被配置为通过修改原有的配置文件中的文件系统加载项来更新存储设备的配置文件。

[0062] 在一个实施例中,以上参考方法100所述的块S105-S106的操作在此同样分别适用于单元305和306,因此不再赘述。

[0063] 如本领域技术人员能够理解的,装置300还可以包括图3中未示出的其它单元;并且在一些实施例中,图3中的某些单元可以被省略。

[0064] 在提出本公开的实施例的方法和装置之前,在存储系统中不存在本地和远程多重数据保护的解决方案。例如,利用VNX Repv2进行的数据保护是非同步的,并且在灾难发生时,在目标存储设备侧可能会存在数据丢失。而对于通过VNX同步复制VDM进行保护的数据,其具有距离限制(例如40km)。因此,只能依靠从例如实验室停电之类的故障中恢复。而如果例如地震或者洪水等大面积的灾难发生,则由于距离较近,两个存储设备站点可能都将遭受灾难影响,并导致两个存储设备均无法工作,从而导致数据服务中断。

[0065] 利用本公开的实施例的方法和装置,能够提供本地和远程复制的共存,例如使VNX中同步复制和异步复制共存。这使得能够提供实时数据保护,也能够实现几百公里的距离

的数据保护。根据一些实施例,异步复制的恢复中能够利用同步复制会话技术来避免初始复制阶段并缩短恢复时间。利用一些实施例,能够同时提供文件级和块级数据保护的综合的解决方案。

[0066] 通过本公开的一些实施例的方法和/装置,用户能够针对VDM创建同步会话,并且针对VDM中加载的文件系统创建Repv2会话。

[0067] 通过本公开的一些实施例的方法和/装置,存储系统能够在灾难发生时提供VDM的同步复制的失效切换,并且在失效切换后恢复FS异步复制会话,从而提供本地和远程多重数据保护。

[0068] 本领域技术人员将容易地认识到,各种上述各种方法中的块或者步骤可以通过编程的计算机来执行。在本公开中,一些实施例还意在涵盖程序存储系统,例如,数字数据存储介质。这是机器或计算机可读的并且编码机器可执行或计算机可执行的指令程序。其中,该指令执行上述方法的一些或所有步骤。程序存储系统可以是,例如,数字存储器、诸如磁盘和磁带的磁存储介质、硬盘驱动器或光学可读数字数据存储介质。该实施例还意在涵盖编程为执行该上述方法的步骤的计算机。一些实施例还意在涵盖一种装置,该装置包括至少一个处理器;以及至少一个包括计算机程序代码的存储器,其中该至少一个存储器和该计算机程序代码被配置为:与该至少一个处理器一起,促使该装置执行方法100。

[0069] 在附图中示出的装置的各种元件的功能,可以通过使用软件、专用硬件以及与适当软件相关联的能够执行软件的硬件、或者固件、或者其结合来提供。当由处理器提供时,该功能可以由单个专用处理器、由单个共享处理器或由多个单独的处理器来提供。此外,术语“处理器”可以包括但不限于,数字信号处理器(DSP)硬件、网络处理器、专用集成电路(ASIC)、现场可编程门阵列(FPGA),用于存储软件的只读存储器(ROM)、随机存取存储器(RAM)和非易失性存储装置。还可以包括其他常规和/或定制的硬件。

[0070] 本领域技术人员应当理解,说明书和附图仅仅说明本公开的实施例的原理。因此,应当理解,本领域的技术人员将能够设计出各种布置,虽然这里没有明确地描述或示出,但是该布置体现本公开的实施例的原理并且被包括在本公开的实施例的精神和范围内。此外,这里阐述的所有示例主要旨在明确仅用于教学目的,以帮助读者理解本公开的实施例的原理和发明人贡献的用于促进本领域的概念,并且应被解释为不限于这些具体阐释的示例和条件。而且,这里阐述本公开的实施例的原理、方面和实施例的所有阐述及其具体示例也意在包含其等同物。

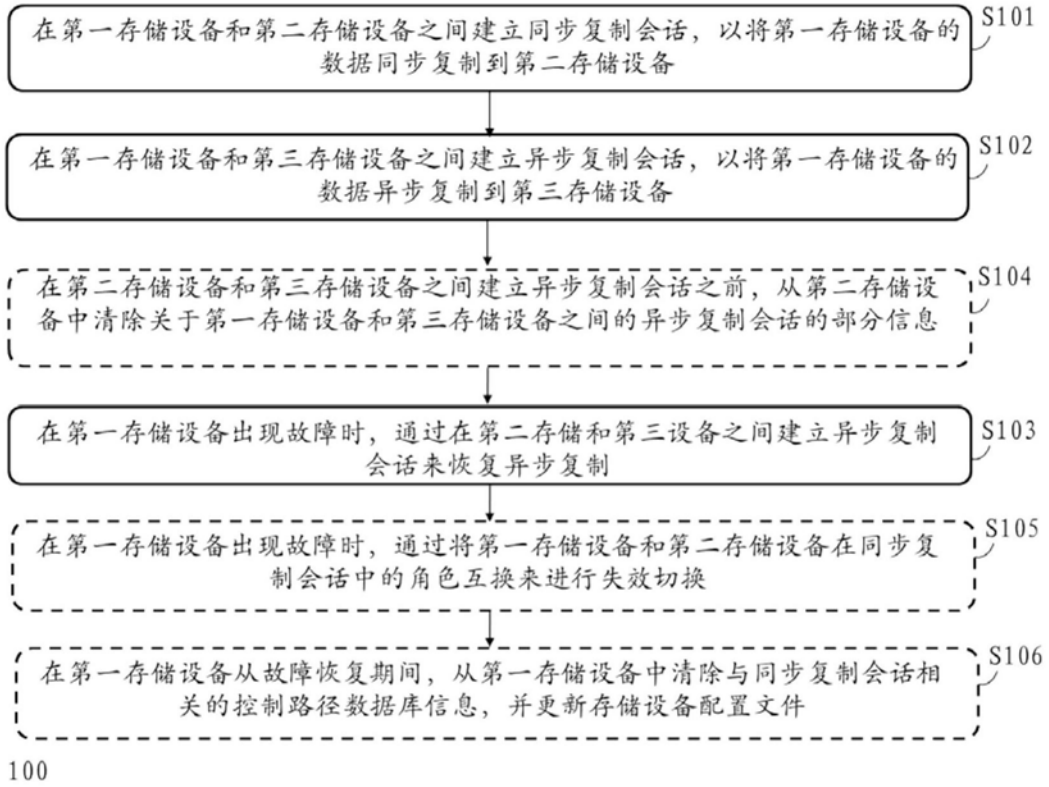


图1

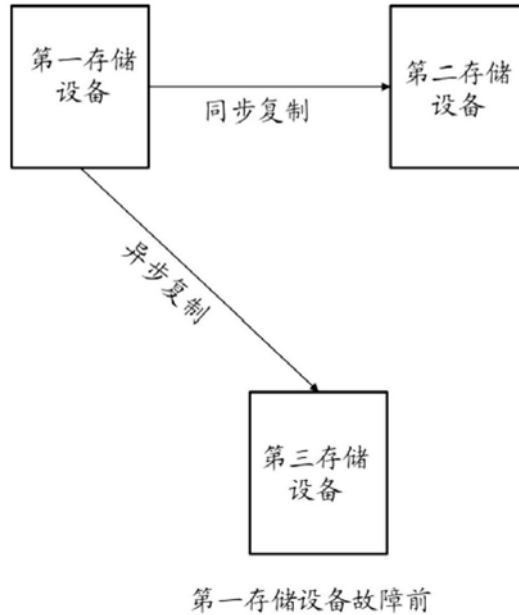


图2a

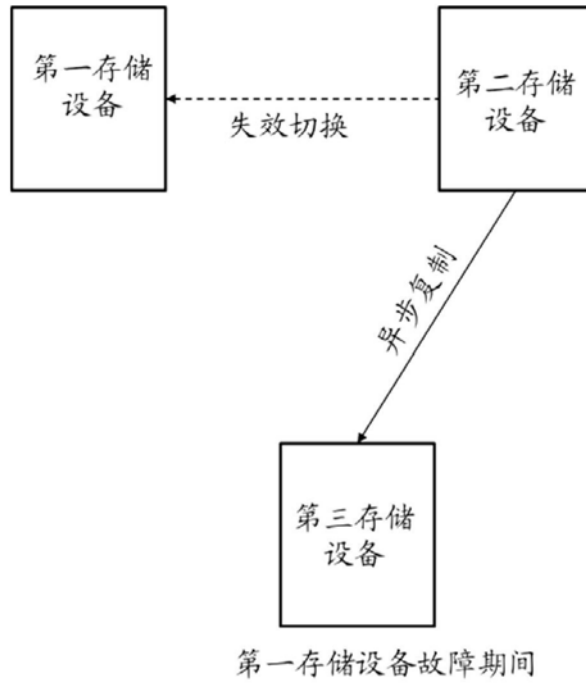


图2b

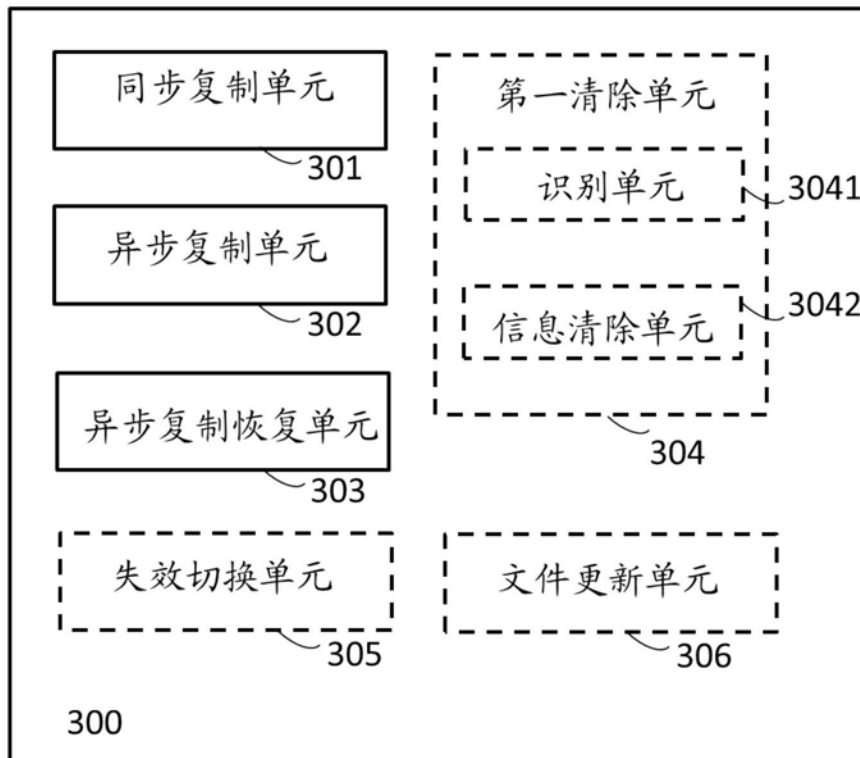


图3